



Reflections on Geographic Variations in U.S. Health Care

Jonathan Skinner, PhD
The Dartmouth Institute for Health Policy and Clinical Practice
Department of Economics, Dartmouth College
Jon.skinner@dartmouth.edu

Elliott S. Fisher, MD, MPH
The Dartmouth Institute for Health Policy and Clinical Practice
Department of Medicine, Dartmouth Medical School
Elliott.fisher@dartmouth.edu

March 31, 2010 (updated May 12, 2010)

Executive Summary

With funding from the National Institutes for Health and not-for-profit foundations, the goal of the Dartmouth research team has been to improve our understanding of the causes and consequences of variation in the way health care is delivered around the country. Why do some primary care physicians order more than twice as many CT scans as their colleagues in the same practice? Why are the rates of coronary stents three times higher in Elyria, Ohio, compared with nearby Cleveland, home of the famous Cleveland Clinic? And most importantly – what do these differences mean for patients? The aim of the research is to advance the science of health care delivery – to understand what actually happens to patients and what can be done to make care better.

During the past several years, the Dartmouth research has drawn attention from policy makers, politicians and stakeholders in health care reform. While many of the findings are broadly accepted and their implications are already being translated into practice, some confusion about the Dartmouth work remains. In this summary, we provide brief responses to the questions that have been raised. A more detailed discussion, with full citations to peer-reviewed journal articles, follows.

The Dartmouth Atlas shows a more than two-fold variation in per capita Medicare spending in different regions of the country. Do these findings take into account differences in the rates per procedure that Medicare pays? We have recently implemented price-adjustments in our Atlas measures. As it turns out, price doesn't matter much, except in New York City which is something of a special case. Adjusting for price differences leads to only a modest decline in overall variations. It is utilization -- the *amount* of care delivered to patients -- that explains most of the regional variation in Medicare spending.

Regions where Medicare spends more have more poor people. Doesn't poverty explain the differences in spending? In a recent *New England Journal of Medicine* paper, using a sample of 15,000 Medicare Beneficiaries, we found that poverty explained little of the variation in health care spending across regions – at most 4 %.

So why do some critics still believe that income matters? It is well known that poverty goes hand in hand with poor health, and we know sicker people generally need more care. Yet people with low income also suffer from limited access to care, so they don't always get the treatments they need.

Do more expensive regions and hospitals have sicker patients? On average, expensive regions have sicker patients, but as we have shown, their higher illness levels explain only fraction of the overall differences in regional variations. A recent study by MedPAC suggested a larger role for illness in explaining regional variations, but their adjustments suffer from well-known biases: people are more likely to be “diagnosed” with a disease when their physician or hospital treats them more intensively. This bias makes patients in high-intensity areas appear sicker than they really are. The most reliable approach to addressing the problem of illness-adjustment – following cohorts of similar patients over time – continues to show more than two-fold variation in utilization across the US.

Dartmouth compares the care of patients in the last two years of life at different hospitals, but doesn't that method fail to take into account the possibility that some hospitals are better at preventing death than others? Two issues are confounded here. The first question is whether end-of-life measures accurately predict how intensively hospitals treat patients with other conditions such as heart attacks. As we show, they do. The second is whether higher intensity hospitals achieve better outcomes. The Dartmouth research that has looked at this question focused on patients with specific conditions, such as hip fractures and heart attacks, and followed them for several years to see how they fared. On average, higher spending was not associated with better outcomes.

Is there any evidence that spending more leads to better outcomes? The key question is: spending more on what? Dartmouth research comparing spending differences across both regions and hospitals found that most of the spending was due to differences in use of the hospital as a site of care (versus, say, hospice, nursing home, or the doctor's office) and to discretionary specialist visits and tests. Higher spending on these services does not appear to offer overall benefits. Other Dartmouth research suggests that hospitals spending more on effective care do in fact get better outcomes.

Other recent studies have found that some higher spending hospitals do have better outcomes along some dimensions. More importantly, many health systems are able to provide high-quality care at low cost, suggesting that we don't *need* to spend more to get better outcomes.

Does Medicare spending track spending in the rest of the health care system? The available evidence suggests that hospitals and regions that provide more care to Medicare patients also provide more for their non-Medicare patients. But in under-age-65 insurance markets, prices per procedure can vary wildly across regions -- so total per-capita regional spending in the under-65 markets may not be closely associated with per-capita Medicare spending in those markets. This suggests we should look for savings in each sector of the health care system.

The Atlas is often cited as a source for the estimate that 30 percent of the nation's spending is unnecessary -- what is the evidence? The Dartmouth approach was to ask how much might be saved if all regions could safely reduce care to the level observed in low spending regions with equal quality; we find estimates ranging from 20-30 percent, but view these as an underestimate given the potential savings even in low cost regions. At least three other groups have come to 30% waste estimates: the New England Healthcare Institute, McKinsey, and Thomson Reuters.

How should reimbursement to regions and hospitals be based on the Atlas work? The Atlas measures of cost are reliable, but should not be used to set payment rates until hospitals and their associated physicians can be organized in ways that allow them to improve patient care. Current Atlas measures do provide useful insights to regional health systems and to individual providers who wish to consider how they might reduce overtreatment, improve care, and curb spending.

The key take-home message is that we believe that there is enormous scope for improving the efficiency and quality of US health care. The Dartmouth research suggests that improvements in both cost and quality can be achieved by supporting new models of payment that reward providers for improving quality, managing capacity wisely, and reducing unnecessary care.



I. INTRODUCTION^a

During the past several years, Dartmouth health care research has drawn attention from policy makers, politicians and stakeholders in health care reform. While many of the findings are broadly accepted and their implications have already been translated into practice, some confusion about the Dartmouth work remains, especially in the public arena, where it can be difficult to briefly summarize a complex and often nuanced body of research. It is therefore worthwhile to provide a brief review of the aims and history of the Dartmouth research, explain where the Dartmouth Atlas Project fits in this work, and then address specific points of confusion.

Dartmouth research – understanding variations and bringing science to practice

For over 30 years faculty at Dartmouth have taken the lead in exploring the questions raised by the well-documented variations in medical practice, health care spending, and patient outcomes that are observed when comparing individual physicians, the populations cared for by hospitals or medical groups, and the populations residing in different regions of the United States. John Wennberg [1] inspired and led this work and recruited a multi-disciplinary faculty to join him at Dartmouth. The importance of these questions has led investigators around the country – and the world – to join the work of improving our understanding of what is actually going on in clinical care.

With funding from the National Institutes for Health and private foundations, the goal has been to improve our understanding of the causes and consequences of these variations. Why do some primary care physicians order more than twice as many CT scans as their colleagues in the same practice? Why are the rates of coronary stents three-fold higher in Elyria, Ohio, compared with nearby Cleveland? Why has per capita Medicare spending in suburban Long Island grown so rapidly that health care costs there account for an extra one billion dollars *annually* compared to the slower growth in San Francisco? [2] And most importantly – what do these differences mean for patients? The aim of the research is to advance the science of health care delivery – to understand what actually happens to patients and what can be done to make care better.

As the work progressed, it became clear that the research had important implications for both patient care and health policy. Two main themes emerged. One, *the patient should play a central role in decision-making*. Research exploring variations in the rates of common surgical procedures, such as prostate surgery, and knee replacement, highlighted the lack of evidence on the risks and benefits of many treatments and the difficulty that patients faced understanding these risks and making well-informed choices. [3] [4] This work suggested the need for methods to help patients more accurately understand risks and to make well-informed choices in consultation with their clinician. [5]

^a We are grateful to Reed Abelson and Gardiner Harris of the *New York Times* who asked us the questions considered in this FAQ. We are also grateful to Shannon Brownlee for valuable comments and suggestions.

The second theme to emerge from the Dartmouth research suggested that *large variations in expenditures are not necessarily associated with better health outcomes*. In other words, spending more does not always lead to better care or better health: it's not how much you spend, it's how you spend it.[6] Per-capita Medicare spending varies by more than two-fold across U.S. regions. Research on the causes of these differences has found that higher spending is due largely to the overall intensity of discretionary care provided to similar patients: how much time they spend in the hospital, how frequently they see physicians, and how many diagnostic tests and minor procedures they receive. [7] The earlier research also found that greater use of these services was generally not associated with better quality of care or better outcomes. [8] [9] [10-11] [12]

Unwarranted variations in practice and spending -- those not due to patients' needs or preferences -- pointed to important opportunities not only to improve the care provided to patients, but also to improve the quality and hence the efficiency of the U.S. health care delivery system.[13] As the work progressed, however, it became clear that few physicians, hospital leaders, policy makers or members of the public were aware of the variations or their implications. The Atlas was established to engage this broader public by shining a light on the variations in practice and spending across all U.S. communities, including their own.

A brief history of the Atlas

The first Dartmouth Atlas of Health Care, edited by John E. Wennberg and Megan Cooper in 1996, described the variations in practice and spending observed across U.S. regions. The goal was to highlight the phenomenon of variations by focusing on the one group with near-universal health insurance coverage – the Medicare population -- for whom comprehensive data could be obtained and analyzed. Early measures included expenditures among all Medicare enrollees, the per-beneficiary supply of physicians, and utilization measures such as back surgery rates, cardiac procedures, ICU days, and other common surgical rates. These continue to be updated and posted on the web, with appendices included with details on methods.^b

A decision was made early on to use the Atlas simply to report the data, adjusted only for age, sex, and race. We understood at the time that there could be other reasons why regions differ with regard to spending, for example by the prevalence of illness, or socioeconomic status. But as we show below, in statistics the cure is often worse than the disease, as “adjusting” for many factors can impose additional biases. The early Atlases were a way to make public and transparent the Medicare claims data – and the questions and insights raised by the variations -- without imposing potentially subjective or imperfect adjustments on the numbers.

The new 2008 Atlas presented a departure from previous Atlases. The idea of studying patients at the end of life across different hospitals originated in earlier work by John Wennberg and colleagues.[14] The idea was that patients near death suffer from similar illnesses whether

^{bb} See for example www.dartmouthatlas.org (click on “download” for the data. Methods can be found at the end of the atlases, which are available also for free download in pdf format. See for example the Appendix in the 1996 Atlas, available at <http://dartmouthatlas.org/atlas/96Atlas.pdf>.

they're in Hawaii or Louisiana, so that comparing how these patients are treated could yield insights about the intensity of care across hospitals. The most recent 2008 Atlas therefore focused on this group, but limited the sample to people with serious chronic illness (thus excluding otherwise healthy people who suddenly die) and adjusted for differences across regions in the prevalence of specific diseases.

Why this new approach for the 2008 Atlas? Over time, we have become more interested in “relative efficiency” and “accountability,” the idea that hospitals and physician practices are more or less efficient and might be held accountable for both the quality and costs of care they provide. Thus we wanted to focus at the level of the hospital, rather than the region. (Rewarding on the basis of regions would not provide sufficient incentives for any one institution or group of physicians to improve.) Our traditional measures, however, could not at that time be used for individual hospitals, since there are many Medicare enrollees with no obvious links or ties to *any* hospital in a given year. Thus focusing only on people who happened to be admitted to hospital X would lead to severe biases: if hospital X admitted lots of healthy people unnecessarily, they would appear to have both low costs (per admission or patient) and good outcomes. For this reason, we focused on this sample of chronically ill Medicare patients near the end of life, because among this group, nearly all patients get their care regularly at a specific hospital.

We continue to explore other dimensions of health care in the Atlas series. For example, a new Atlas, sponsored by the Robert Wood Johnson Foundation, measures quality indicators across regions, including ambulatory-sensitive admissions to hospitals and eye examinations for diabetic patients.^c

2. ARE REGIONAL VARIATIONS REAL?

Whether one looks at variations in the use of evidence-based services (like appropriate testing for diabetic patients), variations in rates of elective surgery (such as cardiac stents), or differences in per-capita spending across regions, the first question that arises is: Are these variations explained by other legitimate factors that also varies across regions or hospitals, such as differences in prices paid per procedure, poverty, or health? We address each in turn.

Variations in Medicare spending; what is the role of price differences?

A key question is whether differences across regions in expenditures reflect real differences in utilization. Another way of putting that question is: *Have differences in spending by region or individual hospital taken into account differences in the prices Medicare pays in different regions or the policy-related payments that are made to hospitals, such as graduate medical education (GME) and disproportionate share hospital (DSH) payments?*

In both our early and more recent work, we have addressed the issue of price adjustments by applying a common “price” or Medicare reimbursement rate to specific procedures and physician visits^d across regions in the U.S.[15] This means that a given surgical procedure in New York

^c <http://www.dartmouthatlas.org/af4q.shtm>

^d Diagnostic related groups (DRGs) and relative value units (RVUs).

City would count just as much towards total spending in the region or at an individual hospital as the same procedure in Enid, Oklahoma. (Note that this approach also removes all GME and DSH payments – a key point to which we return below.) Our price adjustment approach was based on the Medicare Payment Advisory Commission (MedPAC) approach, with additional simplifications on our end. Ultimately, the two approaches are very similar.

The bottom line? We find that price adjustment reduces the degree of regional variation across the U.S., but not by much. For example, after adjusting for price, some rural regions that were formerly middling in terms of spending now look much higher; while other regions, San Francisco for example, which were formerly in the middle of the pack now look to be low-cost. Even after price adjustment, Miami and McAllen Texas are the highest cost regions in the country.

Why then the difference between a recent MedPAC report,^e that seems to suggest much less regional variation than what we find, and our *Health Affairs* paper? The reason is their *illness adjustment* method, a topic we revisit below.

New York is an interesting —and exceptional -- case. A lot of Medicare dollars flow to the New York metropolitan area. But as our adjustment shows, the high rate of spending per enrollee is due in large measure to the very high amount paid per hospital discharge (DRG) – as much as 50 or 60 percent above the U.S. average. This is not solely because prices are higher in New York – after all, prices are high in San Francisco too – but also because New York City hospitals train so many medical residents. One-sixth of all medical residents in the U.S. train in New York hospitals, according to one estimate, and Medicare reimburses hospitals very generously for having large staffs of residents. The region ends up looking expensive overall because the hospitals receive added payments per discharge with their generous DSH and GME payments.

The key question, however, is about utilization. Hospital utilization rates can be decomposed into two components: the admission rate (how likely patients are to be admitted) and the length of stay (how long they remain as an inpatient during a given admission). In most regions of the country, differences in hospital utilization are driven by the admission rate. But because length of stay is much longer than average in NYC than elsewhere, we see that the overall number of hospital inpatient days in NYC hospitals is at the upper end of the U.S. range. One possible explanation is that some of the additional payments for graduate medical education help pay for longer lengths of stay in NY hospitals.

Should New York City be judged a high-utilization region? In terms of Medicare DRGs and hospital admissions, it is still above average, but is no longer an outlier. But in terms of actual time spent in the hospital, patients with at least 10 different doctors, total physician visits (and total dollars per patient) it ranks near the top of the country.

Variations in Utilization: What is the role of poverty?

^e http://www.medpac.gov/documents/Dec09_RegionalVariation_report.pdf

Some critics of the Dartmouth work have been concerned that regional variations simply reflect different socioeconomic patterns across areas. *Have differences in spending by region or individual hospital been adequately adjusted for the socioeconomic status of the patients?*

As noted above, the first-round Medicare expenditure measures were simply adjusted by age, sex, and race. In part, this was because it's not always clear how one should "adjust" for income. Suppose for example that higher income people consume more health care not because they are sicker, but just because they demand more and get more. It's not clear that one should adjust for income in this case, because one might view high income people commanding more health care resources as "unwarranted" and thus worthy of documentation.

Even so, income explains very little of the variation across individuals[16] or regions in health care spending. This latter point was shown most recently in a paper where we used a sample of 15,000 people in the Medicare Current Beneficiary Survey, with individual information about their income, and their health status. [17] Briefly, we found that only a trivial fraction of spending across high- and low-cost regions could be explained by differences in income.

So why do some critics still believe that income matters? It is perhaps understandable that Figure 1, reprinted from our paper[17] might cause casual readers to think that poverty could explain differences in spending across regions or hospitals, since the graph shows substantially more health care spending among those in lower income groups.

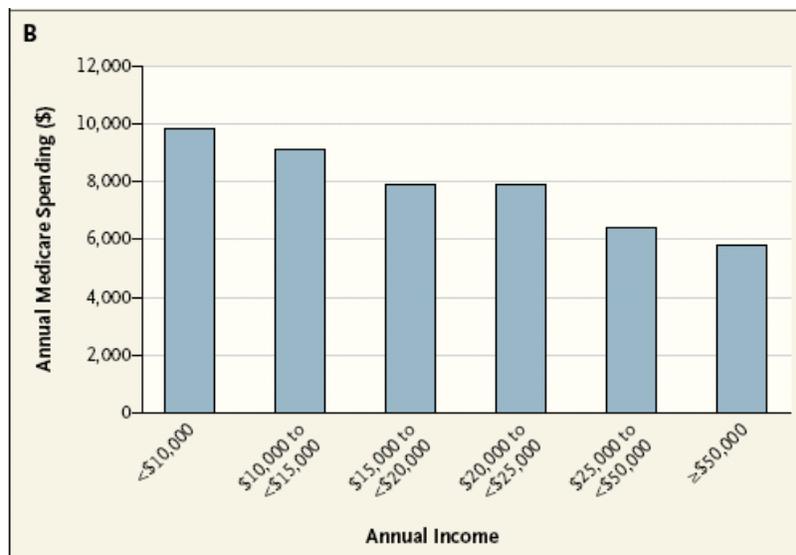


Figure 1: The Association Between Individual Income and Individual Medicare Expenditures

Source: Sutherland J, Fisher E, Skinner J, NEJM 2009.

This comparison does not control for other important factors explaining Medicare spending, for example that low income people tend to be older, and older people account for more Medicare spending. But let's use these numbers anyway, since even our harshest critic, Richard Cooper, has embraced them.

It is true that regions where Medicare spends more have more elderly people who live in poverty. Indeed, Cooper even shows a strong correlation between the *percentage* of people in the region in poverty and the level of spending in the region.^f Unfortunately, his comparison doesn't answer the question we need to know – *how much* of the spending differences across regions can be explained by poverty?

So let's use the Medicare expenditure data in Figure 1 to make a much simpler and intuitive assessment of the importance of poverty in explaining differences across regions in spending. Figure 2 shows just how much poorer are the elderly people who live in the highest-cost 20th percentile of the country (Quintile 5) -- those who live in regions like Miami and Los Angeles -- compared to the people who live in regions in the lowest cost 20th percentile of the country (Quintile 1), like San Jose or Albany, NY.

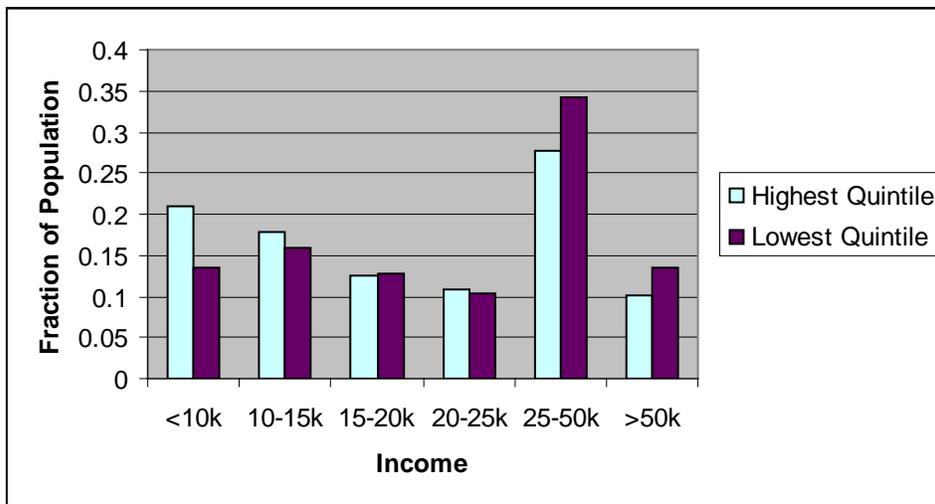


Figure 2: The Distribution of Income in the Medicare Population by Quintile of Health Care Expenditures

Source: Medicare Current Beneficiary Survey, author's calculations.

While the high cost (highest quintile) regions do have more people living in poverty, the important thing to note from Figure 2 is that poverty is prevalent beyond big cities; even Green Bay WI has a poverty rate of 16.7%.^g The data from Figure 2, combined with the income-based spending estimates in Figure 1, can tell us how much of the total variation in spending can possibly be explained by poverty and the poor health associated with poverty.^h And the answer

^f <http://buzcooper.com/2009/11/23/dartmouth-loses-another-battle-against-the-poor/>

^g <http://www.city-data.com/poverty/poverty-Green-Bay-Wisconsin.html>

^h The actual frequencies are, starting from < 10K and going up to \$50K+, .210, .179, .125, .110, .276, .101, for the high cost region and .135, .158, .127, .103, .341, .136 for the low cost region. To construct "predicted" spending based on income, multiply each of the spending measures in Figure 1 (\$9841, 9086, 7921, 6398, 5814) by their relevant frequency above to determine how the different income distributions affect Medicare spending.

is: not much. Spending in the high-cost regions is predicted to be \$343 greater than in the lowest-cost regions (\$7898 versus \$7555) because of income differences. But the real difference in spending is \$3300. That means that a mere 4 percent of average Medicare spending, a small fraction of the total spending gap, is explained by differences in the level of poverty among regions. A newer study finds an even smaller association between income and Medicare expenditures.[18] This is because income cuts two ways; it may increase health needs, but it is also well known to reduce access to health care.

Now in fairness to our critics, it is well known that poverty goes hand in hand with poor health. So perhaps what they are really saying is that sickness explains the difference in spending, because sicker people need more health care. Below, we demonstrate that health differences across regions *are* important and do need to be considered in comparing Medicare spending across regions, but health differences do not explain away regional variations.

Variations in Utilization: the role of poor health and the challenge of risk adjustment

Another question raised in recent years is the following: *Have differences in spending by region or individual hospital been adequately adjusted for the differences in the severity of illness?*

One of the ways we have studied this question involved using the 15,000 Medicare enrollees in the Medicare Current Beneficiary Survey (MCBS) mentioned above, which also includes data on self-reported health and disease prevalence (e.g., smoking, diabetes, obesity). [17] Race is relatively unimportant in explaining spending, and income, as we showed above, is not far behind. Of the three, health status is by far the most important factor explaining differences across regions, explaining about \$600 of the total \$3300 gap in spending. In other words, some regions are indeed sicker than others.

While a 15,000-person study may seem enormous, it does not allow us to make any inferences about the role illness plays in determining the level of spending or utilization in any particular region. This is an important point – while we know that, on average, billions of dollars in spending variations cannot be explained by illness or other factors, it is more difficult to pinpoint exactly how much of Detroit’s higher spending (for example) is explained by higher levels of underlying illness.

The MedPAC analysis of regional spending tried to get at this question using a different method than we did. They used measures of illness derived directly from the Medicare claims data, rather than self-reported illness. Their results suggested that differences in spending across regions was far less pronounced than our research suggested. For example, McAllen, Texas – an outlier in our data -- no longer appeared to be particularly expensive, because patients there appeared to be sicker than in other regions.

The question is, did they do the risk adjustment correctly? Price-adjusted health care spending in McAllen was nearly \$14,000 per Medicare enrollee in 2006, compared to \$8,167 on average. Are patients in McAllen, Texas, really that much sicker than patients in lower-spending regions?

The MedPAC analysis used a measure developed by Medicare to adjust payments to Medicare Advantage plans: hierarchical condition category scores, (HCCs) which counts the number of different diagnoses that patients have received over the course of a year. Some of the diagnoses are themselves inferred from whether the patient had a specific procedure. This method assumes that the more diagnoses a patient has, the sicker they are, with further adjustment for exactly what the diagnoses are: patients diagnosed with an ear infection, say, are not as sick as those diagnosed with heart disease.

The problem with HCC scores are that they depend upon what doctors do. Physicians who are aggressive in screening their patients for heart disease, say, will naturally find more people who have heart disease. And we know that one of the characteristics of high cost regions and hospitals is they tend to do more testing than in lower cost locations. This creates circularity – regions that have doctors who do more testing will have patients with more diagnoses and thus will appear to have sicker patients. This was shown in a *New England Journal of Medicine* study that first demonstrated Medicare enrollees who moved to a high-intensity region looked just the same as enrollees who moved to a lower-cost region prior to a move. [19] However, following the move, the people who moved to the high-intensity regions were diagnosed with more diseases than their counterparts who moved to lower-intensity regions. In other words, McAllen is not necessarily sicker than other lower-spending regions – indeed, by many measures it looks remarkably healthy among the over-65 population – but because far more people there have been told by their physician that they have a disease.

If risk-adjustment is so tricky, then what makes us think that our end-of-life measure used in the 2008 Atlas is any more accurate?

We think end-of-life utilization measures for the chronically ill represent at least a good starting point for measuring hospital utilization and intensity. Of course, the problem with end of life measures is that decedents may differ across regions with regard to their illnesses, thus biases such measures.[20] Say there are a lot of lung cancer patients in one region. People who die of lung cancer tend to die more quickly than, say, people with heart failure. If you look back two years from death, the average lung cancer patient may have been relatively healthy for the early part of the two years, then sickened and died quickly, thus costing less than the average heart failure patient who is quite ill over the entire two years.

But the 2008 Atlas measure has already addressed this concern by adjusting for the type of disease people have, and whether they have multiple diseases. (Although we are not able to adjust for different types of cancer.) In other words, a hospital serving a higher fraction of heart failure patients than lung cancer patients would not be penalized when using our methods.ⁱ

We have also begun to explore other methods to risk adjust patients at individual hospitals that are not subject to the bias present in the MedPAC approach. One of the most promising methods

ⁱ Suppose for example that hospital X has 66 lung cancer patients and 34 heart failure patients, while hospital Y has 66 heart failure patients and 34 lung cancer patients. Since overall there are 100 heart failure patients and 100 lung cancer patients, the adjusted spending measure for each hospital would be equal to .5 times the average cost of their heart failure patients plus .5 times the average cost of their lung cancer patients.

we've used is to focus on cohorts of patients with similar diseases. For example, nearly every hip fracture patient or heart attack patient is admitted to the hospital, and so comparing what happens to heart attack or hip fracture patients across regions or hospitals represents one way to measure risk-adjusted differences in spending.

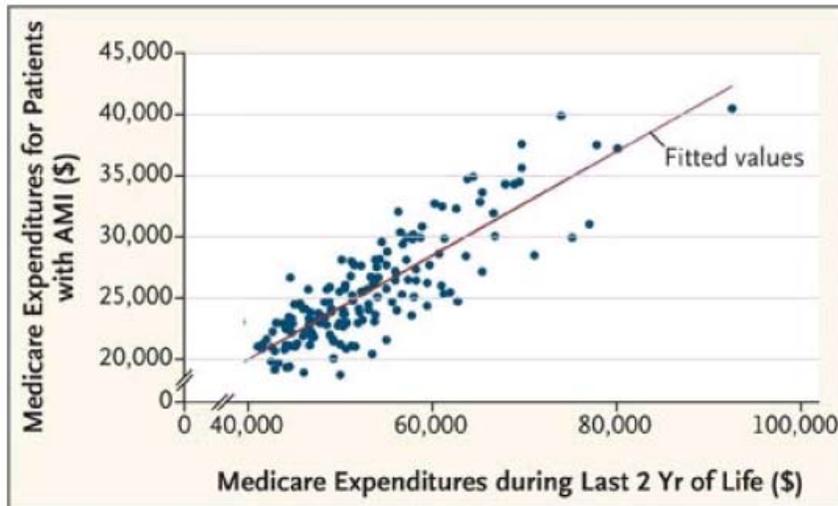


Figure 3: Hospital-Specific Spending on Patients with Heart Attacks (AMI) and End of Life Care, for 144 Large Hospitals. and Source: Skinner, Staiger, Fisher, NEJM 2010.

Figure 3 (from our 2010 *New England Journal of Medicine* paper) compares risk-adjusted cost measures for this “look-forward” approach using heart attack patients, with the 2008 Atlas measures for end-of-life spending. Each dot represents one of the 144 largest US hospitals, and shows total inpatient spending for heart attack patients (on the vertical axis) and spending on patients in the last two years of life (on the horizontal axis). For the heart attack data, we adjusted for age, sex, race, income and poverty rates in the patient’s zip code, anatomical location of the heart attack (some heart attacks are more serious than others), other diseases present at admission like cancer, dementia, and diabetes, and price adjustment for the hospital referral region. Thus we have done a careful job of risk adjustment. Yet notice two things. First, plenty of variation remains even in risk-adjusted spending for heart attack patients across hospitals: from under \$20,000 to more than \$40,000. Two, spending on risk-adjusted heart attack patients is tightly correlated with our end-of-life measure of spending (the correlation coefficient is 0.85) – in other words, it doesn’t matter how one measures costs, pretty much the same result is obtained. Perhaps an average across these two (or more) measures could be used to measure hospital spending. In other words, we’re not far from accurate spending measures at the hospital level particularly if we include other cohorts such as hip fracture or cancer patients.

3. IS HIGHER SPENDING ALWAYS BETTER? IT DEPENDS WHAT YOU SPEND IT ON

This is the most important question of all. If spending more (and doing more) results in better outcomes, then perhaps places where spending is lower need to do more for their patients. On the other hand, if spending more does not produce better outcomes, that suggests that a great deal of utilization in higher-spending regions and hospitals represents wasted effort and money – not to

mention unnecessary exposure to the risk of medical intervention for patients. So the critical question is less about the overall association between spending and outcomes (whether positive, negative or uncorrelated), but about what the additional spending is for and how these services might – or might not – contribute to better outcomes.

On the end of life care, the controversy seems to be whether the benefits of more intensive care are ignored in our research because there is no way to measure outcomes other than death. How does our research capture the benefits of more intensive care for specific diseases or patients?

There have been a large number of studies by the Dartmouth group showing that more spending doesn't necessarily lead to better health. In the most comprehensive study using national data and millions of people, Elliott Fisher, David Wennberg and coauthors looked at what happened to patients who were hospitalized with one of three different conditions: heart attack, hip fracture, and colon cancer (this last category included only those who were in the hospital for surgery). [9] [10] First they divided regions into five equally sized quintiles based on average end-of-life spending in the region. In other words, the "look-back" end-of-life measures were used to assign regions to quintiles; Los Angeles was a high-cost region (for end-of-life care) and so anyone with a heart attack, hip fracture or colon cancer in Los Angeles was assigned to that quintile.

They then "looked forward" to see what happened to each of the three risk-adjusted cohorts of patients, starting on the day the patient was hospitalized for her heart attack, hip fracture, or colon surgery. The most important conclusion to be drawn from this paper was that patients treated in regions where utilization (and spending) were higher received about 60% more care over the first year after their initial hospitalization but did not experience better risk-adjusted outcomes when compared with those treated in regions where utilization was lower. In other words, more was not necessarily better. It's important to note that their results do not depend on end-of-life measures – they tried a number of different approaches to measuring regional utilization, such as one-year spending for hip fracture or heart attack patient, and they got the same results: more spending and more utilization did not yield better outcomes. (See the results shown in the Appendix of those papers). The other important finding was that the additional spending in the high cost regions was largely devoted to hospital stays, physician visits, specialist consults and tests that were not ordered by the physicians in the low cost regions.

While the results of this paper are quite robust, end-of-life measures of intensity do not always imply that more is worse. Joseph Doyle Jr. at the MIT Sloan School of Management used end-of-life measures to sort hospitals into high and low cost intensity regions.[21] He found that tourists admitted to an emergency room in a high-cost region survived longer than tourists admitted to a low-cost region ER. Clearly in some cases more can be better, particularly for acute illnesses among tourists with acute illness presenting at the ER. He found no association between spending more and better outcomes, however, for the much larger group of Florida residents admitted through the ER.

The Dartmouth researchers and others have used several methods to get at this question of whether more spending and utilization results in better outcomes. Some of the most recent studies include:

- Jonathan Skinner and colleagues looked at changes over time in spending and outcomes for heart attack patients. [12] (No end-of-life patients were used in this study!) They found that increased spending and utilization over time did not lead to better outcomes.
- In another study, Skinner and Douglas Staiger found that for heart attack patients, spending more could result in better outcomes – especially for effective low-cost care such as beta blockers and aspirin.[22] In other words, it matters more how the money is spent, and much less how much is spent.
- Amber Barnato and colleagues examine the association between greater intensity of care (higher utilization) and 30-day and 6-month survival in Pennsylvania hospitals.[23] They found for elderly patients higher intensity of care was associated with improved survival at 30 days, but the improvements waned by 6 months. Barnato's measure of intensity focuses on the relative rate at which each hospital used the following potentially life-sustaining treatments: ICU use, mechanical ventilation, hemodialysis, tracheostomy and feeding tubes. The importance of this article lies not only in the finding that seriously ill patients are likely to live slightly longer with more intensive care. The innovation is that they attempt to clarify what *type* of care might offer additional benefits. They also found that the benefits from greater intensity are found primarily in those hospitals with relatively low intensity to start with, hospitals that may lack even a fully-staffed ICU or rapid response teams. And the article noted that the magnitude of the association is small; a 1.5 percent improvement in the chance of surviving an extra six months is associated with thousands of dollars in additional spending and invasive procedures.
- A paper by Lena Chen and colleagues looked at a cross-section of U.S. hospitals discharging Medicare patients with congestive heart failure (CHF) and pneumonia.[24] Hospitals in the highest-cost quartile had slightly higher quality of care measures for CHF patients (89.9% versus 85.5%) and lower mortality (9.8% versus 10.8%), but the converse was true for pneumonia patients: 85.7% versus 86.6% for quality measures, and 11.7% versus 10.0% for mortality. On net, this appears to be a wash; there is little consistent association between spending and outcomes in this study.
- Finally, a recent study by Michael Ong and colleagues found a very strong association between more spending and better outcomes for congestive heart failure (CHF) patients in six California hospitals.[25] Thus the result might appear to show that more is, in fact, better.^j

Because the Ong paper has received so much attention in the media, we reviewed their findings in a bit more detail. We turned to our Medicare claims data set of 1.1 million heart attack patients to compare one-year spending and mortality for the six hospitals in the Ong et al. sample. Using heart attack patients, we came up with much the same results as the Ong et al. study did for congestive heart failure – a strong and significant positive correlation between spending and health outcomes – *for the 6 hospitals in the study*. But for the larger sample of all hospitals in California, there is no apparent association, and for the entire United States, the

^j The major focus of their paper is to contrast their “look forward” measures of intensity with the “look back” measures used in the *Annals of Internal Medicine* study mentioned above. It turns out that for the 6-hospital sample, the Ong et al. result, that “more is better,” is *also* invariant to whether one uses “look forward” or “look back” end-of-life measures of intensity. Ranking their 6 hospitals by end-of-life spending demonstrates a similar very strong negative association between end-of-life spending and mortality, with a p-value of 0.07.

correlation coefficient between mortality and expenditures is flipped around, so more appears worse.

How can these puzzling results be reconciled? Figure 4 shows a hypothetical scatter plot between expenditures and mortality. The red dots represent the (hypothetical) sample of 6 hospitals, and there is a clear negative association between spending and lower mortality, meaning more is better. However, the entire sample of hospitals (red and blue dots combined) shows a positive association between expenditures and mortality. In other words, one must be wary of generalizing from a small sample of hospitals.

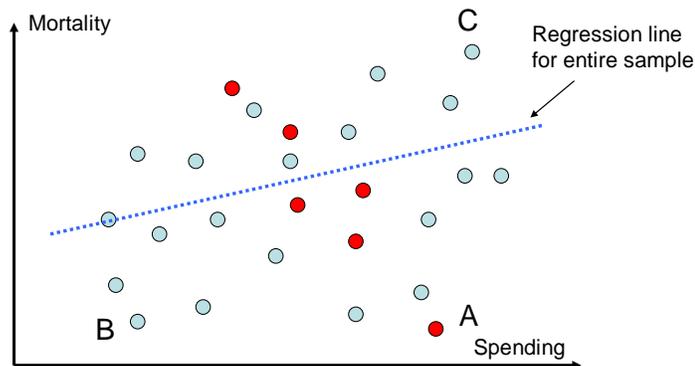


Figure 4: Hypothetical Graph Showing the Correlation between Hospital Spending and Quality for a Subgroup of Six Hospitals (Red Dots) and All Hospitals (Red and Blue Dots)

Does this mean that more spending really leads to more people dying? Or equivalently, that by slashing spending in high-cost regions, we'd end up saving lives? Probably not. Look again at Figure 4 – it looks like what happens when you shoot buckshot at the side of a barn. Some hospitals may get excellent results with high rates of Medicare spending (like the red dot labeled A). But other hospitals get good results without Medicare spending nearly as much (the dot labeled B). More worrisome are the hospitals that cost a lot and get poor results, labeled C. As we discuss below, there is more to inefficiency than just measuring differences along the horizontal axis – some hospitals cost more than others to treat identical patients. Just as worrisome is variation along the vertical axis – at some hospitals, the quality of care is very poor, meaning that patients die at much higher rates. Both excess costs *and* excess deaths contribute to inefficiency in the U.S. health care system, a point to which we return below.

A reasonable conclusion of these studies would be that there are large variations in spending that are largely unrelated to outcomes, and because there are both regions and hospitals that provide both excellent outcomes and lower costs, we have an important opportunity to learn from them – and to improve the overall efficiency of U.S. health care.

How big are the potential savings?

The Atlas is often cited as a source for the estimate that 30 percent of the nation's spending is unnecessary and therefore can be reduced without affecting the quality of care being delivered. What is the evidence?

Our approach has been to define regions with low utilization as benchmarks, and show that higher utilization does not, on average, produce better outcomes. Thus spending in excess of the benchmark that does not appear to yield health benefits is deemed “waste.” The 30 percent figure is mentioned in the *Annals of Internal Medicine* (2003) paper, which in turn cited a 1997 *National Tax Journal* paper. In this 1997 paper, we compared what spending would have been had we attained the illness-adjusted and price-adjusted per capita spending levels of a low cost region such as Richmond (\$3842) or Minneapolis (\$3722) compared to the average price-adjusted per capita spending in the U.S. (\$4878). Choosing Minneapolis as the benchmark for a relatively efficient region implied that U.S. spending was 31% higher than what spending might have been under Minneapolis-style health care.

At the time the *Annals* paper was submitted, another paper used a simpler benchmark against the low-cost regions, and in this case estimated waste to be 29 percent of average Medicare expenditures.[13]

Other estimates from different papers are as low as 20 percent -- particularly in recent years where previously low-cost providers have adopted some of the expensive habits of the higher cost regions. But more recently, the New England Healthcare Institute has expanded the measurement of waste beyond just one dimension and estimated inefficiency at about 30 percent of total spending.[26] Similarly, the McKinsey Report estimates that the U.S. spends \$650 billion more than spending (30% of the total) that would have occurred in health care systems in other developed countries, without clear evidence of a benefit,[27] while Thomson Reuters suggests 33% waste.[28]

We think that the 30 percent estimate could be too low, for two reasons. First, our relative efficiency measures assume that all hospitals and physicians in low-cost regions are perfectly efficient. We don't capture the waste in, for example, elective surgery that patients would not have wanted had they been fully informed, which may amount to as much as 25 percent of elective surgeries. This point was also made by Melinda Beeuwkes Buntin and David Cutler, who similarly argue that waste exceeds 30 percent.^k

And second, we are not capturing the inefficiency arising from poor health outcomes. Consider for example the study by Baicker and Chandra [8] who compared state-level Medicare expenditures with state-level process quality measures (like beta blocker use after heart attacks, or flu shots) published by Jencks et al. [29] As is shown in Figure 5 below, they found a negative association between spending and quality. Clearly there appears to be waste in terms of dollar spending – but more to the point, there is also inefficiency arising from remarkably low rates of effective care. For example, the Jencks study shows that in 2000-2001, the use of beta blockers within 24 hours of a heart attack – a low-cost treatment shown to reduce mortality by as

^k http://www.americanprogress.org/issues/2009/06/pdf/2trillion_solution.pdf

much as 25% -- ranged from 50% (Arkansas) to 86% (NH).¹ Placing a dollar value on the resulting loss of lives arising from the underuse of effective treatments would imply inefficiency measures in U.S. health care substantially above 30%.

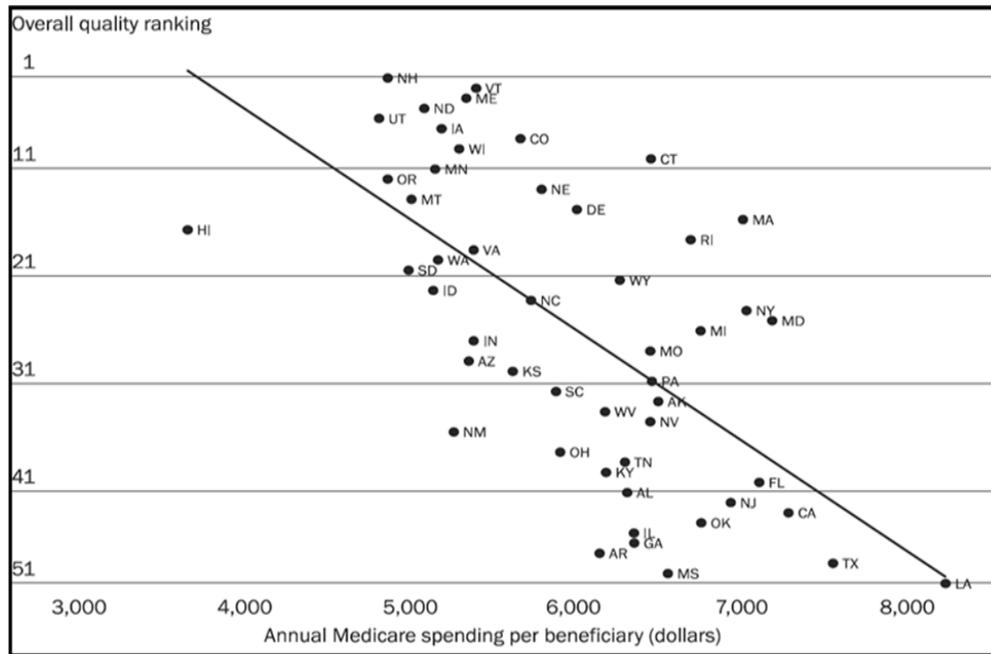


Figure 5: The Association Between Annual Medicare Spending Per Enrollee and Overall Quality Ranking of Health Care. Source: Baicker and Chandra, Health Affairs, 2004.

In sum, it is hard to imagine anyone arguing that that total waste, broadly defined, is less than 30 percent. The challenge, of course, is whether we can actually reduce that waste through health care reform.

4. FROM INSIGHT TO ACTION: IMPLICATIONS FOR POLICY AND PRACTICE

A number of questions have arisen over the past year or two about how the insights from variations research can be applied in the context of health care reform. Here we consider several questions – how well can the Dartmouth results for Medicare data be extended beyond that population, what do these measures say about the quality of care for patients, and how should the payment structure be reformed in light of these findings.

Are Dartmouth Results True for Anybody Besides Medicare Patients?

¹ Since the time of the study (2000-2001 data), these rates have risen to near 100 percent among most hospitals. As well, we recognize that these data in Figure 5 are relative rankings – this is why we consider actual measures of beta blocker use in the text.

One of the questions of increasing interest is the following: *Does Medicare spending track overall health care spending and what research supports the idea that the most expensive doctors and hospitals in Medicare are the most expensive in the private sector?*

This is a great question, and we are slowly learning more about the under-65 population and how complicated and opaque local markets for private insurance can be. A few observations:

(a) The price per procedure is known to vary tremendously across regions and hospitals depending on the market power of the insurance company versus the hospitals. What this means is that the cost per patient in the under-65 population may bear little relation to the cost per patient in the over-65 population solely because the prices paid in the under-65 population are all over the map, while prices in the over-65 Medicare market are administered by CMS and don't vary that much. There are a variety of theories about how hospitals might be shifting costs from the Medicare market to the private market and vice versa, but these are still somewhat speculative. A recent paper suggests that when hospitals feel pressure to constrain their costs they are able to do so.[30]

(b) Earlier work from California has shown a reasonable correlation between utilization for the over 65 Medicare population, those covered by Medicare Advantage plans, and the under 65 population. [31] The recent study by Mike Chernew and others showed a correlation between over and under 65 hospital utilization rates [32]. Finally, new evidence from Amitabh Chandra at Harvard is shown below in Figure 6: The correlation between Part B (physician) spending for the elderly, and a very different measure of intensity – the rate of Cesarean Sections among normal birthweight babies. These are obviously two very different markets! Yet there is a marked correlation ($0.47, p < .01$), with McAllen, Texas, showing one of the highest rates of Cesarean Sections in the country. This suggests that the way doctors and hospitals deliver care is similar in any given region, regardless of who is paying.

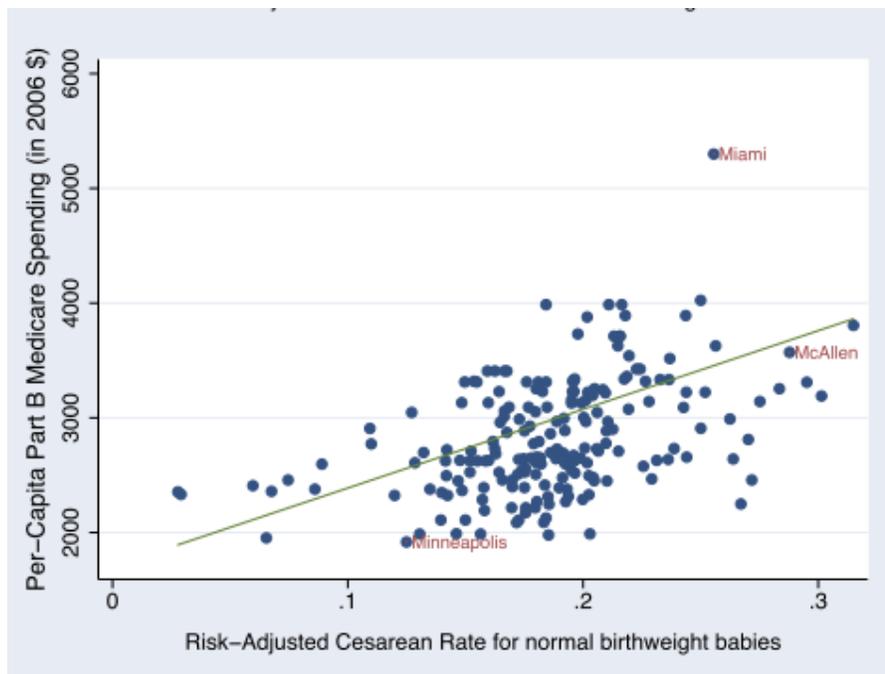


Figure 6: The Correlation Between Part B Expenditures and Risk-Adjusted Cesarean Section Rates for Normal Birthweight Babies

(c) We have been looking at health care expenditures for the under-65 Medicare population, most of whom are on Social Security Disability Insurance. They currently comprise 20% of total Medicare enrollees (and about the same fraction of spending), yet the correlation appears remarkably high between the under-65 and over-65 populations.

(d) Suppose one accepts for the moment a critic's view that Medicare spending is not predictive of under-65 spending. This does not reduce importance of variability in health care spending overall, since estimates of under-65 spending show as much variability as in the over-65 population. In other words, if there is as much variation in the younger population as in the older population – and it appears there is – then there is as much potential savings from the younger population as the Medicare population.

More importantly, the key threat to the financial security of the U.S. government is the growth in Medicare spending. So even if the Medicare data were used solely to improve incentives and outcomes in the Medicare program, this would be a huge leap forward towards both improving the financial stability of the federal government, and improving the quality of care for Medicare enrollees. And we would anticipate substantial spillover effects to the under-65 markets as well.

What does the Dartmouth Atlas say about the quality of care, especially at the end of life?

We know that patients in the last months and even years of life receive very different amounts and kinds of care, depending upon where they live and the hospitals they choose. Patients loyal to New York University Medical Center, to take an example at the high end of utilization, spend on average 54 days in the hospital in the last two years of their lives, and receive 110 inpatient physician visits, many of which involve invasive tests and procedures. At the low end of the spectrum, patients who are loyal to, say, Stanford University Medical Center, spend 20 days in the hospital in the last two years of life and receive 56 doctor visits.

Some of the care dying patients receive at these two hospitals is clearly effective and evidence based, and that kind of care is a great quality measure, for example a prescription for beta blockers for heart attacks. An example of a negative quality measure, a measure that indicates poor quality, is giving patients with dementia a feeding tube – which clinicians agree is a bad thing and hence a marker for poor care. A recent JAMA article by Joan Teno at Brown University and colleagues documented the use of feeding tubes per hospital admission across the United States. [33] They showed that demented patients who are admitted to the ICU of a high-spending hospital are 2.6 times more likely to have a feeding tube than similar patients admitted to a low-spending hospital. That is, one of the risks of going to a high-spending hospital is being treated with very unpleasant and ineffective care.

Teno and her colleagues provided an appendix with some specific measures for individual hospitals. Average feeding tube placement in the U.S. was about 6.4 percent in 2006-07 for dementia patients admitted to the hospital. For Minnesota hospitals in their study, average rates

of feeding tube use were less than 1 percent. For Florida, they were much higher – as much as 37 percent of dementia patients in one hospital. Hospitals can also be highly idiosyncratic in their use of treatments. For instance, Cedars-Sinai and UCLA Medical Center, two prominent academic medical centers in Los Angeles, rank high on the use of ICU days for patients in their last 6 months of life, but 18 percent of Cedars-Sinai patients with dementia were reported to have had a feeding tube, compared to 0 percent at UCLA.

What is the value to patients and their families of this intensive care? Again, Joan Teno and her team at Brown[34] have shown using interviews of survivors that people dying in institutions have much worse experiences leading up to death than those dying in hospices or at home.

Can the Atlas be used to identify institutions and regions that waste money on unnecessary care?

The Dartmouth Atlas currently provides a wide variety of measures for each of the 306 Hospital Referral Regions (based on place of residence) and hospitals (based on where patients received their inpatient care). These include measures of per-beneficiary spending, utilization rates for surgical procedures and other categories of care (such as physician visits), and some measures of quality that are summarized from either Medicare claims data or surveys carried out by Medicare.

Given what we know so far about the overall lack of correlation between utilization and outcomes or quality, higher spending regions and hospitals are likely to be overusing, or over-delivering services as inpatient stays, physician visits, and tests. Put another way, there is good evidence to suggest that these regions and hospitals could provide equal or better care at lower costs if they eliminated unneeded discretionary services. But this does not mean that with our current measures, we can point to a specific high-cost hospital and conclude that they are providing below-average quality.

How should reimbursement to regions and hospitals be based on the Atlas work?

The research (our own and others) suggests that there are real opportunities to improve care and outcomes while slowing the growth in spending. It also suggests that the drivers of overuse include: fee-for-service payment, lack of accountability for overall costs and quality, the local supply of hospital resources (beds, ICU beds, specialists), outpatient facilities and physicians, and local culture of medical practice. Our work has thus contributed to the understanding of the kinds of reforms that enjoy widespread support, including promoting more integrated, organized care, better performance measures to support accountability, and new payment models.

There are critics, of course. Peter Bach, for instance, voiced one objection to using Dartmouth data in a recent Perspectives piece in the *New England Journal of Medicine*, in which he claimed that hospitals that spend more and save lives in the process will somehow be penalized because they will appear to be inefficient.[20] This is actually not the case. First, hospitals that spend lots of money on their potential end-of-life patients will look less expensive if they are successful

in saving lives.^m And more importantly, if the money was spent well, the expensive hospitals would end up receiving top marks on quality, because their patients with heart attacks or hip fractures would flourish, and their readmission rates would remain low). So in this scenario, the expensive hospital could still be rewarded for its high quality care. Of course this raises the question of whether they really need all that money to save the additional lives, particularly if other hospitals are providing equal quality and outcomes at lower cost.

As noted above, we've calculated end-of-life spending and look-forward expenditures for heart attack patients by hospitals, and they are highly correlated and seem pretty reliable. So at least for Medicare expenditures, we think we're close to getting good measures of costs.ⁿ Since much of the challenge in health care reform is how to reimburse hospitals for Medicare costs, these would be a good place to start in terms of measuring costs – assuming that there is some kind of “accountable care” organization in charge that can exert real control over how these patients are being treated.

Quality measures are still being created and vetted in practice, and there exists potential for better measures that include (for example) biomarker measures from patients to measure how well their blood pressure or cholesterol is being controlled. These quality measures are not always highly correlated with one another, and may reflect specific clinical groups; the geriatrics group in hospital X may be better than in hospital Y, but the cardiology group at hospital X could be worse. So clearly progress towards “industrial strength” reliable quality measures is still required.

That said, we don't think these measures should be broadly implemented – yet – to impose costs on hospitals or physician groups, simply because there are so few hospital/physician groups that can control what's done for their patients. (And so here we agree with Dr. Bach.) This limitation is of course a troubling commentary on the state of U.S. health care, that there is so much fragmentation and so little accountability for patients. We have also argued against the so-called value index -- an effort to adjust provider payments based on regional spending. [35] [36]

So how might a hospital/physician group gain the ability to control their own costs and given the opportunity and incentives to improve? We view Accountable Care Organizations (ACOs) as a necessary building block that should be implemented before going ahead with rewards or penalties to hospital/physician groups. In fact, the ACO pilots that we are supporting now are using total-per-capita costs as the spending measure for setting targets, with work under way to figure out the best way to adjust for illness differences.

^m To see this, suppose that there were 100 chronically ill patients being treated at a high-intensity hospital, where treatment runs \$90,000 on average, or 50% above the average spending of \$60,000. If the hospital saves 3 heart attack patients -- patients who would have otherwise died in the average hospital -- and it cost \$180,000 per life saved, then spending at that hospital would be somewhat less than 50% above the average, because they had managed to move 3 very expensive patients out of their “end of life” cohort.

ⁿ The critical point, however, is that we are open to revision and are continually striving to improve our measures. For example, leaders at Cedars-Sinai hospital objected to our approach to measuring intensive care unit beds, which lumped step down and high-intensity beds. In our next version, the Atlas team revised the measures as the Cedars-Sinai team suggested so that readers could distinguish these two levels of ICU beds.

At the same time, we believe that the Atlas measures provide useful insights to regional health systems and to individual providers who wish to consider how they might do under new payment models and who wish now to look for opportunities to improve their efficiency.

How does our work support the concept of Accountable Care Organizations?

There are many arguments for moving toward more integrated delivery systems that are accountable for their performance. Our work has emphasized the role of local capacity, or supply of medical resources, and discretionary decision-making in a fee-for-service system as important drivers of spending and spending growth. [7] [37] Our work has also shown that most physicians practice in natural referral networks around one or more hospitals [38] and that some systems provided equal or better care at lower costs. [7] [12] These empirical findings and many others (such as work by Steve Shortell and colleagues) show that integrated systems are better able to improve care provides further rationale for ACOs. [39] These strands of research came together in the policy proposals developed by MedPAC and ourselves to develop payment models that would foster integration and accountability for quality and overall costs (and thus capacity). [40] The research suggested that improvements in both cost and quality could be achieved by supporting new models of care that reward providers for improving quality, managing capacity wisely, and reducing unnecessary care. [40]

5. CONCLUSIONS

The key take-home message is that we believe that there is enormous scope for improving the efficiency and quality of US health care. For any type of meaningful health care reform, however, we must be able to measure both costs and health outcomes accurately. Thus identifying high- and low- cost institutions is a first step towards reducing unwarranted variation in expenditures, thereby saving money. And identifying high- and low-quality institutions can improve the safety and reliability of care, as well as ensuring that patients get the treatment they value by supporting informed patient choice.

References

1. Fleming, C., et al., *A decision analysis of alternative treatment strategies for clinically localized prostate cancer. Prostate Patient Outcomes Research Team.* JAMA, 1993. **269**(20): p. 2650-8.
2. Fisher, E.S., J.P. Bynum, and J.S. Skinner, *Slowing the growth of health care costs-- lessons from regional variation.* N Engl J Med, 2009. **360**(9): p. 849-52.
3. Barry, M.J., et al., *Watchful waiting vs immediate transurethral resection for symptomatic prostatism. The importance of patients' preferences.* JAMA, 1988. **259**(20): p. 3010-7.
4. Fowler, F.J., Jr., et al., *Symptom status and quality of life following prostatectomy.* JAMA, 1988. **259**(20): p. 3018-22.
5. O'Connor, A.M., et al., *Toward the 'tipping point': decision aids and informed patient choice.* Health Aff (Millwood), 2007. **26**(3): p. 716-25.
6. Weinstein, M.C. and J.A. Skinner, *Comparative effectiveness and health care spending-- implications for reform.* N Engl J Med. **362**(5): p. 460-5.
7. Fisher, E.S., et al., *Variations in the longitudinal efficiency of academic medical centers.* Health Aff (Millwood), 2004. **Suppl Web Exclusive**: p. VAR19-32.

8. Baicker, K. and A. Chandra, *Medicare Spending, The Physician Workforce, And Beneficiaries' Quality Of Care*. Health Aff (Millwood), 2004.
9. Fisher, E.S., et al., *The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care*. Ann Intern Med, 2003. **138**(4): p. 273-87.
10. Fisher, E.S., et al., *The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care*. Ann Intern Med, 2003. **138**(4): p. 288-98.
11. Sirovich, B.E., et al., *Regional variations in health care intensity and physician perceptions of quality of care*. Ann Intern Med, 2006. **144**(9): p. 641-9.
12. Skinner, J.S., D.O. Staiger, and E.S. Fisher, *Is technological change in medicine always worth it? The case of acute myocardial infarction*. Health Aff (Millwood), 2006. **25**(2): p. w34-47.
13. Wennberg, J.E., E.S. Fisher, and J.S. Skinner, *Geography and the debate over Medicare reform*. Health Aff (Millwood), 2002. **Supp Web Exclusives**: p. W96-114.
14. Wennberg, J.E., et al., *Use of hospitals, physician visits, and hospice care during last six months of life among cohorts loyal to highly respected hospitals in the United States*. BMJ, 2004. **328**(7440): p. 607.
15. Gottlieb, D.J., et al., *Prices don't drive regional Medicare spending variations*. Health Aff (Millwood), 2010. **29**(3): p. 537-43.
16. McClellan, M.B. and J. Skinner, *The incidence of Medicare*. Journal of Public Economics, 2006. **90**: p. 257-276.
17. Sutherland, J.M., E.S. Fisher, and J.S. Skinner, *Getting past denial--the high cost of health care in the United States*. N Engl J Med, 2009. **361**(13): p. 1227-30.
18. Zuckerman, S., et al., *Clarifying Sources of Geographic Differences in Medicare Spending*. New England Journal of Medicine, 2010.
19. Song, Y., et al., *Regional Variations in Diagnostic Practices*. New England Journal of Medicine, 2010.
20. Bach, P.B., *A map to bad policy--hospital efficiency measures in the Dartmouth Atlas*. N Engl J Med, 2010. **362**(7): p. 569-73; discussion p 574.
21. Doyle, J. *Returns to Local-Area Healthcare Spending: Using Shocks to Patients far from Home*. 2010; Available from: http://www.mit.edu/~jjdoyle/doyle_vacation_feb2010.pdf.
22. Skinner, J. and D. Staiger, *Technology Diffusion and Productivity Growth in Health Care*, in *Working Paper Series*. 2009, National Bureau of Economic Research: Cambridge MA.
23. Barnato, A.E., et al., *Is Survival Better at Hospitals With Higher "End-of-Life" Treatment Intensity?* Medical Care, 2010. **48**(2): p. 125-132.
24. Chen, L.M., et al., *Hospital cost of care, quality of care, and readmission rates: penny wise and pound foolish?* Arch Intern Med. **170**(4): p. 340-6.
25. Ong, M.K., et al., *Looking forward, looking back: assessing variations in hospital resource use and outcomes for elderly patients with heart failure*. Circ Cardiovasc Qual Outcomes, 2009. **2**(6): p. 548-57.
26. NewEnglandHealthcareInstitute, *Waste and Inefficiency in the U.S. Health Care System*. 2008.
27. McKinsey, Global, and Institute, *Accounting for the Cost of US Health Care: A New Look at Why Americans Spend More*. 2008.
28. Kelly, R. *WASTE IN THE U.S. HEALTHCARE SYSTEM PEGGED AT \$700 BILLION IN REPORT FROM THOMSON REUTERS*. 2009; Available from: http://thomsonreuters.com/content/press_room/tsh/waste_US_healthcare_system.

29. Jencks, S.F., E.D. Huff, and T. Cuerdon, *Change in the quality of care delivered to Medicare beneficiaries, 1998-1999 to 2000-2001*. JAMA, 2003. **289**(3): p. 305-12.
30. Stensland, J., Z.R. Gaumer, and M.E. Miller, *Private-Payer Profits Can Induce Negative Medicare Margins*. Health Aff (Millwood).
31. Baker, L.C., E.S. Fisher, and J.E. Wennberg, *Variations in hospital resource use for medicare and privately insured populations in California*. Health Aff (Millwood), 2008. **27**(2): p. w123-34.
32. Chernew, M.E., et al., *Geographic correlation between large-firm commercial spending and Medicare spending*. Am J Manag Care, 2010. **16**(2): p. 131-8.
33. Teno, J.M., et al., *Hospital characteristics associated with feeding tube placement in nursing home residents with advanced cognitive impairment*. JAMA. **303**(6): p. 544-50.
34. Teno, J.M., et al., *Family perspectives on end-of-life care at the last place of care*. JAMA, 2004. **291**(1): p. 88-93.
35. Pardes, H. and E.S. Fisher (2010) *Getting to Yes on Health Care Reform*. The Huffington Post, DOI: February 24, 2010.
36. Skinner, J., D. Staiger, and E.S. Fisher, *Looking back, moving forward*. N Engl J Med, 2010. **362**(7): p. 569-74; discussion 574.
37. Sirovich, B., et al., *Discretionary decision making by primary care physicians and the cost of U.S. Health care*. Health Aff (Millwood), 2008. **27**(3): p. 813-23.
38. Fisher, E.S., et al., *Creating accountable care organizations: the extended hospital medical staff*. Health Aff (Millwood), 2007. **26**(1): p. w44-57.
39. Shortell, S.M. and L.P. Casalino, *Health care reform requires accountable care systems*. Jama, 2008. **300**(1): p. 95-7.
40. Fisher, E.S., et al., *Fostering Accountable Health Care: Moving Forward In Medicare*. Health Aff (Millwood), 2009.